

Config Watch #5

September 9, 2026 · MarketMania Research · Weekly series

WEEKLY · CONFIG WATCH · Issue #5 · the ceiling rises, the replays do not · Issue date September 9, 2026 · Language English · Methodology v1.1 + Config Watch amendments (approved 12.08, hash e66c7e8c864a2233)
Search windows: **WK** Aug 31-Sep 7, 2026 and **M30** Aug 10-Sep 7, 2026 -- both frozen and Monday-aligned (WK = Mon Aug 31 00:00 -> Sun Sep 6 23:59 UTC; M30 = Mon Aug 10 00:00 -> Sun Sep 6 23:59 UTC, 4 full weeks), sharing the cut-off edge Mon Sep 7 00:00 UTC, so slots dated Sep 7 fall outside both. Slots fall inside the window; trades settle past its right edge at their horizon -- the same 2-3 day luft the engine has used since issue #1 and that the live presets run on.

RESEARCH SNAPSHOT

16,398 simulations, 0 errors, in two frozen collects	WK Aug 31-Sep 7, 2026	M30 Aug 10-Sep 7, 2026	7-stage search: scan -> grid -> refine -> stake sweep -> per-ticker -> OOS replays
engine 1.1; \$1,000 deposit, 10x leverage	stake is universe-aware from sim #1 (table below)	maturity gate: ≥ 15 closed (WK) / ≥ 20 (M30)	criteria frozen in writing before launch (w228 plan)
walk-forward: 34 past cards + 4 live presets replayed verbatim on both windows	6 of 34 past cards Survive (net-positive on both windows)	stability: 0 configs across WK and M30 (night run)	models: 7 (grok axis = grok-4.6 (incl. 4.5-era))
best WK council +\$510.4 (+51.0%)	best M30 council +\$1,717.1 (+171.7%)	zero-trade cells: 996 (weekly) / 1,711 (night run)	margin-skip alerts on leaders: 1 (WK) / 0 (M30) rows

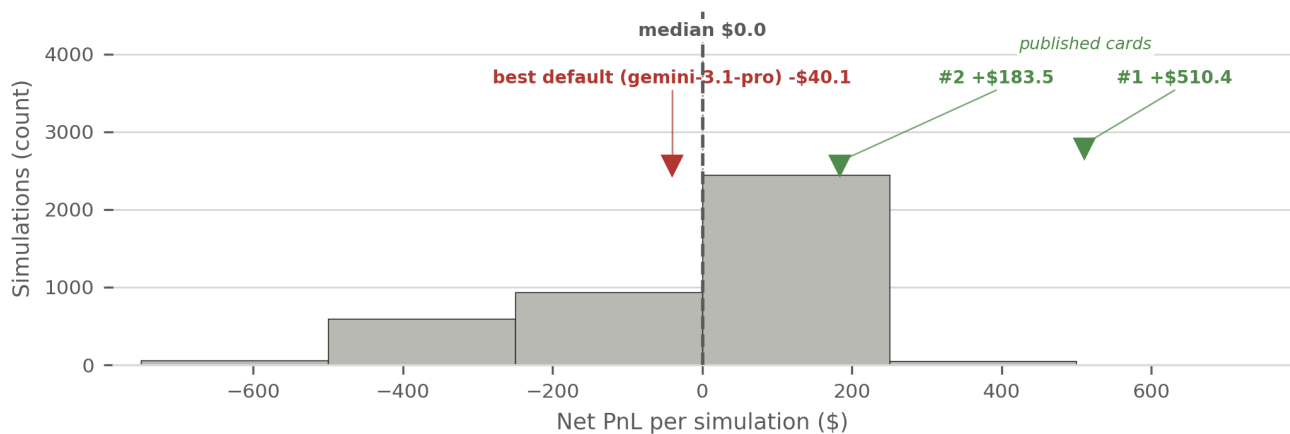
KEY FINDING · OBSERVATION (the ceiling rises, the replays do not)

The ceiling of the fresh search rose from +\$399.4 to +\$510.4 on the week while only 8 of the 34 cards this series has published are net-positive on that same week, and under the standing two-window rule 6 of the 34 Survive: 24 clear the month, 8 clear the week.

TL;DR

- OBSERVATION -- The standing board grew to 34 cards and thinned.** Issue #4 kept 7 of its 26 past cards under the same rule; this issue replays **34 cards** -- issues #1, #2, #3, #4 and Monthly Config Watch #1 -- on two fresh windows, and **6 of the 34 Survive** (net-positive on both). On the week alone 8 of 34 are positive, on the four-week window 24 of 34. Issue #4's own week champion, replayed verbatim, makes -\$70.0 on the week and +\$480.7 on the month.
- The search ceiling rose on the week.** The best mature council on WK makes +\$510.4 (+51.0% of a \$1,000 deposit, 90.9% win rate, 0.0% maxDD, 66 closed) against issue #4's +\$399.4; on M30 it makes +\$1,717.1 (+171.7%, 80.0% win rate, 28.0% maxDD, 30 closed) against issue #4's +\$1,676.0. The grids behind them: median \$0.0 with 34.6% positive across 4,080 WK cells (1 of them above +\$500), median \$0.0 with 47.6% positive across 4,176 M30 cells (579 above +\$500).
- Calibration and the universe, window by window.** Across the whole stage-2 grid calibration improves 57.8% of the 1,403 matched WK pairs (mean +\$136.7) and 37.8% of the 1,441 matched M30 pairs (mean -\$65.3); pooled, 47.7% of 2,844. On the cards, **both published WK councils are calibrated cells and the one M30 card is not a calibrated cell.** The mature net top-10 is a single universe on each window: SOL+XRP on WK (10 of 10 slots), SOL+XRP on M30 (10 of 10).
- The stability board is empty.** The night collect computes WK and M30 in one pass, so this issue's cross-window board is measured on the two windows the issue prints -- and **0 configurations** are in the mature net top-40 of both. Issue #3's board, the last one measured across the same pair, held 21. The hand-tuned anchor keeps its standing measurement against the whole mature field of each window -- 315 of 1,674 mature WK configs and 370 of 3,252 mature M30 configs beat it on win rate, maxDD and smoothness at the same time. Net PnL stays the headline board (continuity with issues #1-#4).

Stage-2 systematic grid -- net PnL distribution -- WK window (Aug 31-Sep 7, 2026) -- n=4,080



Net PnL across all 4,080 stage-2 grid sims, WK window (Aug 31-Sep 7, 2026), universe-aware stake throughout. Grid median **\$0.0**, 34.6% positive (issue #4's week: median \$0.0, 31.0% positive; its month: \$0.0, 47.7% positive). The spike at \$0 is mostly cells whose entry filters produced no trades (992 zero-trade sims on this window). Range -\$642.6 to +\$510.4; 1 cells finished above +\$500. Red marker = best solo default at table stake; green = the published cards. The same board definition one week earlier put its top-5 at +\$399.4 down to +\$372.2; this week's runs +\$510.4 down to +\$465.9 -- a property of the week, not evidence about either issue's cards. The M30 histogram appears after the M30 cards. Source: grid_hist_wk.json.

Why it matters

MarketMania publishes default LLM-council trading configs. This series asks one narrower question every week: how much apparent performance can large-scale in-sample search extract -- and how much of it survives out-of-sample? Issue #1 set the in-sample baseline, issue #2 delivered the first walk-forward verdict (brutal), issue #3 the second (the opposite), issue #4 the third on two windows at once. Issue #5 delivers the fourth, and it is the first issue whose board carries every generation the series has produced: 34 cards from four weekly issues and the first monthly one, each replayed verbatim on the same two fresh windows. That is the point of the exercise -- a rule that is applied to a growing, never-pruned list is the only walk-forward number that cannot be chosen after the fact.

How we searched

Two frozen collects, run to the same plan with the criteria frozen in writing *before* launch (w228 plan) and printed into the meta of every JSON deliverable: the weekly collect over the WK window, and the night collect, which sweeps the M30 window and the same WK week together in one pass. Each stage below prints **weekly + night** and both totals are printed whole; no calendar-month window is computed this issue. **Every WK number in this issue comes from the weekly collect and every M30 number from the night collect** -- the night run's own second pass over the WK week is used for one thing only, the cross-window stability board, and no number from it is printed beside a weekly-collect number.

- **1. Scan (s1)** -- 476 + 952 sims across council compositions and coarse settings.
- **2. Systematic grid (s2)** -- 4,080 + 8,352 sims sweeping the declared axes (archetype, FH/TF, entry filters, universe, council size, membership, TP/SL source, ladder, break-even, TP/SL shift, calibrate and stake) -- the denominator for every distribution claim below.
- **3. Refinement: ladder + stake sweep (s2b+s3b)** -- 116 + 230 sims around the grid winners.
- **4. Preset neighbourhood grid (s2p)** -- 53 + 106 sims: one-knob neighbours of the live presets that have a local grid this issue.
- **5. Per-ticker probes (s5t)** -- 70 + 110 sims: best finalists split onto single coins.
- **6. Dedicated per-ticker (s6t+s6tc)** -- 580 + 1,115 sims: a compacted independent per-coin grid plus calibrate twins (the standing branch from issue #2).
- **7. OOS replays (s4cw4+s4cw4o+s4cwm1+s4cwm1o+s4cw3+s4cw3o+s4cw2+s4cw2o+s4cw1+s4pre)** -- 60 + 98 sims: every published card and all four live presets, verbatim, on every fresh window, plus origin re-runs of the issue-4 and Monthly-1 cards for the drift baseline.

Total 16,398 simulations (5,435 weekly + 10,963 night), 0 errors, 2,707 zero-trade cells (996 + 1,711); stage sums reconcile exactly in each collect (counts.by_stage). The axis grid was NOT widened relative to issue #4 (anti-overfit rule): the replay branch grew by two cohorts, the search space did not. Equity series, grid histograms, baselines and the calibrate pack are produced inside the same two collects -- there is no separate same-day extras pass this issue.

Stake: universe-aware from simulation #1 (\$1,000 deposit)

Instruments in universe	Stake cap	Sweep values also simulated
1 coin	\$450	\$300 / \$375
2 coins	\$450	\$300 / \$375
3 coins	\$300	\$225 / \$375
4 coins	\$250	\$200 / \$300
5 coins	\$200	\$150 / \$250

A \$1,000 deposit cannot fund five concurrent \$450 positions. Stake is a function of how many instruments the config trades, recomputed on every universe change and swept as its own axis; any simulated stake above the cap for its universe size is marked **concentrated** and excluded from the headline boards (it stays in the open data). Every card also prints **skipped_nofunds** -- entries the engine could not fund. Margin-skip alert rows on the leaders this issue publishes: 1 on WK (weekly collect) and 0 on M30 (night collect); the night collect raised 1 alert in total, the rest on its own second run of the WK week, which this issue does not publish. Source: criteria in both collects.

Models under test (exact engine IDs): claude-fable-5, claude-opus-5, gpt-5.6-sol, deepseek-v4-pro, gemini-3.1-pro, qwen-3.8-max, grok-4.6. The grok axis runs as **grok-4.6 (incl. 4.5-era)** -- one lineage; published grok-4.5 configs replayed out-of-sample get the same 4.5 -> 4.6 lineage remap ([remap] rows).

TOP councils -- WK window (Aug 31-Sep 7, 2026)

#	Council	Coins	FH/TF	Net	Net %	maxDD	Trades	WR
#1	gpt, gemini	SOL+XRP	4h/4h,1h	+\$510.4	+51.0%	0.0%	66	90.9%
#2	fable, opus, gpt, gemini, grok	BTC+ETH	4h/4h	+\$183.5	+18.4%	0.1%	23	82.6%
#	Ladder	BE	Shifts	Conf filter	Source	Entry	Positional	Stake
#1	50:40,100:60	on	TP -10 / SL +75	>=60 (each)	farthest	3/0	reenter	\$450
#2	50:40,100:60	on	TP -10 / SL +75	>=60 (each)	farthest	3/0	reenter	\$450

Against the 4,080-cell stage-2 grid on this window: card #1 lands in the +\$500 to +\$750 bucket (1 of 4,080 cells) and card #2 in the \$0 to +\$250 bucket (2448 cells); 1 cells in the whole grid finished above +\$500 (grid median \$0.0). Card percentiles exported this issue cover the raw net-board leaders and the replay rows, not the mature cards -- no percentile is claimed for the two cards themselves. Net % = net PnL as a percentage of the fixed \$1,000 starting deposit; every simulation runs at **10x leverage** (liquidation modeled by the engine). Entry = max_sideways / max_diff_side.

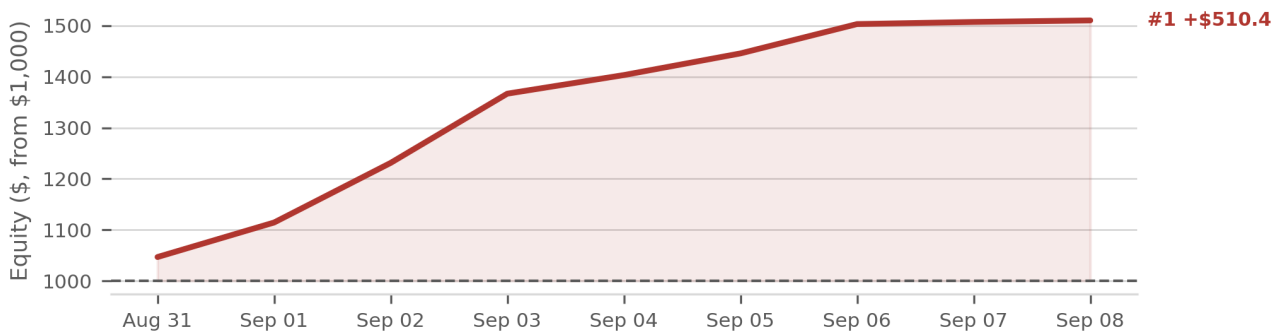
Only two councils are published on this window. All ten slots of the mature net board are the same 2-coin universe (SOL+XRP), so the pairwise-different-universes rule admits exactly one; #2 is the best mature multi-coin council outside that universe on the quality boards -- it comes off the **Smoothness** board (+\$183.5 on BTC+ETH, R2 0.963, ulcer 0.03, 0.1% maxDD, 0 unfunded entries) -- disclosed rather than padded. Read the cards as two universes deep, not three.

On the calibrate axis both cards are calibrated cells -- the second issue running in which every published WK council is a calibrated cell. Card #1 runs the 2-coin table stake of \$450, so the same knobs at \$375 (the sweep value) are on the quality boards at +\$425.3. Neither card carries an unfunded entry.

OPEN IN SANDBOX marketmania.ai/s/cw5-wk1 marketmania.ai/s/cw5-wk2

Each link opens this exact frozen window; results visible without sign-in (embedded share signature). The codes are minted at publish time and are NOT live in this draft; every card's full frozen query signature is in the open-data JSON next to its simulation index.

Published cards equity -- WK (Aug 31-Sep 7, 2026) -- \$1,000 at 10x -- 66 settled trades



Daily settled equity for card #1 above (\$1,000 start, 10x leverage, frozen window; #1 red -- fixed rank colors across all issues). A curve running past the window end is an open-at-cutoff trade settling at its horizon (the 2-3 day left); the curve's last point reconciles to its config's net to the cent. There is no contrast line on this window: the net-board maximum is the same simulation as card #1 (i=2008, 66 closed, above the ≥ 15 maturity gate), so a second curve would be the same curve. Card #2 has no daily series in this issue's equity pack, and the solo and smoothness series stay in the open data (equity.cards) -- the PDF carries one equity chart per window. Source: equity_wk.json.

TOP councils -- M30 window (Aug 10-Sep 7, 2026)

#	Council	Coins	FH/TF	Net	Net %	maxDD	Trades	WR
#1	fable, opus, gpt, gemini	SOL+XRP	1d/4h	+\$1,717.1	+171.7%	28.0%	30	80.0%
#	Ladder	BE	Shifts	Conf filter	Source	Entry	Positional	Stake
#1	50:40,100:60	on	TP -10 / SL +75	≥ 60 (each)	farthest	3/0	reenter	\$450

Against the 4,176-cell stage-2 grid on this window: card #1 is the same simulation as the raw net-board leader (i=3236), so its exported percentile is the card's -- above **99.9%** of the grid. Grid median \$0.0, 579 of the 4,176 cells above +\$500. Net % = net PnL as a percentage of the fixed \$1,000 starting deposit; every simulation runs at **10x leverage** (liquidation modeled by the engine). Entry = max_sideways / max_diff_side.

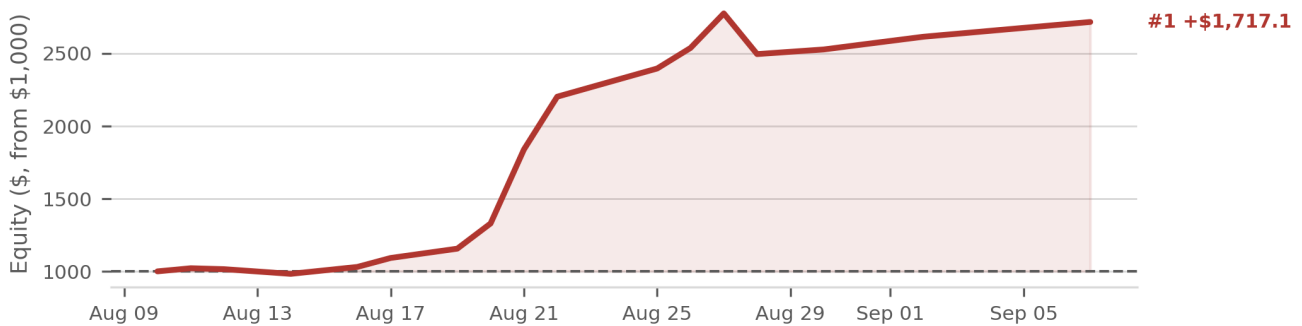
One council is published on this window, not two, and it is the same universe the week collapsed onto. All ten slots of the M30 mature net board are SOL+XRP, and so is every multi-coin council row of all four quality boards (68 of 68) -- there is no second universe anywhere on this window for the pairwise-different-universes rule to admit, so the rule admits exactly one card and the issue publishes one rather than padding the block. The two single-coin rows the boards do carry (SOL) are excluded by the ≥ 2 -coin rule. Read the M30 block as one universe deep.

On the calibrate axis the one M30 card is not a calibrated cell, where both published WK councils are calibrated cells -- the same axis, read on two windows (Calibrate axis below). Card #1 pays for its +\$1,717.1 with a **28.0% drawdown** on 30 closed trades and 0 unfunded entries. The M30 mature pool is 3,252 configs against 1,674 on WK, because four weeks clear the ≥ 20 -trade gate that one week does not.

OPEN IN SANDBOX marketmania.ai/s/cw5-m301

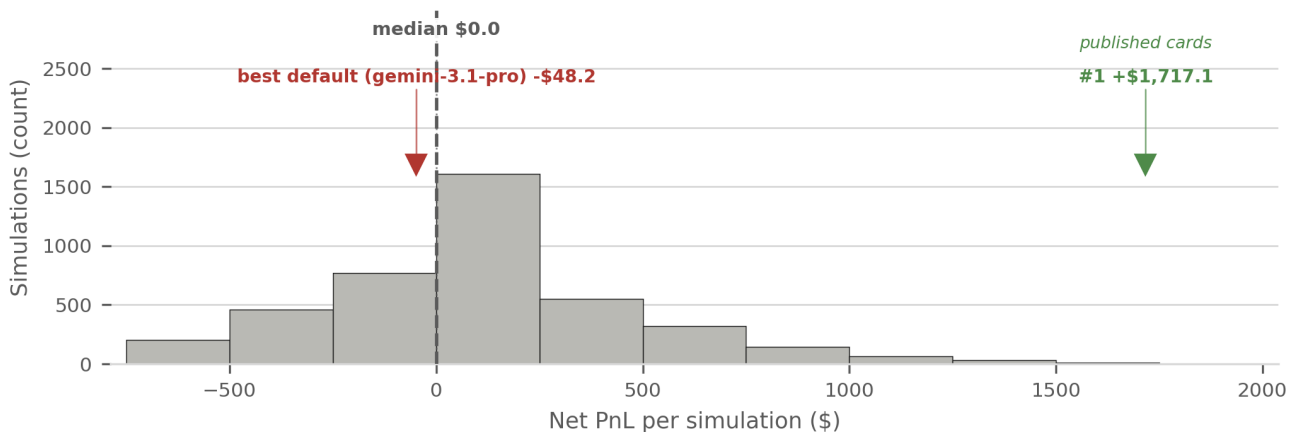
Each link opens this exact frozen window; results visible without sign-in (embedded share signature). The codes are minted at publish time and are NOT live in this draft; every card's full frozen query signature is in the open-data JSON next to its simulation index.

Published councils equity -- M30 (Aug 10-Sep 7, 2026) -- \$1,000 at 10x -- 30 settled trades



Daily settled equity for card #1 above (\$1,000 start, 10x leverage, frozen window; #1 red -- fixed rank colors across all issues). A curve running past the window end is an open-at-cutoff trade settling at its horizon (the 2-3 day luft); its last point reconciles to the card's net to the cent. There is no contrast line on this window: the net-board maximum is the same simulation as card #1 (i=3236, 30 closed, above the ≥ 20 maturity gate). The solo, Calmar, WR and smoothness series stay in the open data (equity.cards) -- the PDF carries one equity chart per window. Source: equity_night.json.

Stage-2 systematic grid -- net PnL distribution -- M30 window (Aug 10-Sep 7, 2026) -- n=4,176



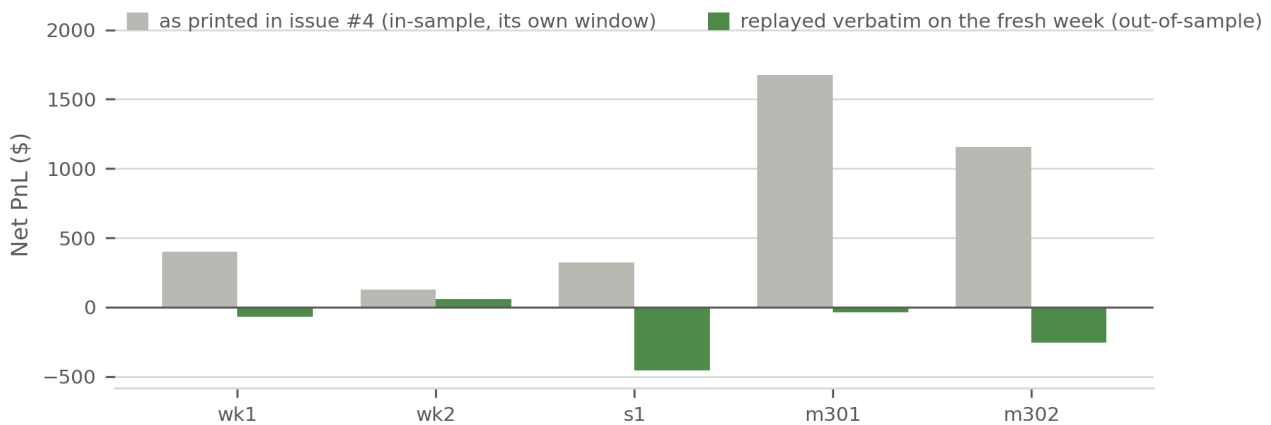
Net PnL across all 4,176 stage-2 grid sims, M30 window (Aug 10-Sep 7, 2026), universe-aware stake throughout. Grid median **\$0.0**, 47.6% positive against 34.6% on the week (issue #4's month: median \$0.0, 47.7% positive). The spike at \$0 is mostly cells whose entry filters produced no trades (752 zero-trade sims on this window). Range -\$707.9 to +\$1,793.9; 579 cells finished above +\$500, against 1 on the week. Red marker = best solo default at table stake; green = the published card. Source: grid_hist_night.json.

Walk-forward: the standing OOS verdict

Every config this series has ever published -- all 34 cards from issues #1, #2, #3, #4 and Monthly Config Watch #1 -- replayed VERBATIM (same knobs, same stake as printed, no re-tuning) on both fresh windows, plus the four hand-tuned presets running live on the replay, listed first. The monthly issue's three cards are judged by the weekly rule, like every other row. The replays and their origin re-runs cost 60 simulations in the weekly collect and 98 in the night collect, which replays every row on both of its own windows.

Verdict rule, unchanged since issue #2. Survived = net-positive on BOTH fresh windows; **Faded** = net-negative on at least one. The live presets and the monthly cards are judged by the same rule. New WK is pure out-of-sample for every row -- the week begins exactly at issue #4's right edge (Aug 31). New M30 (Aug 10-Sep 7, 2026) shares 21 of its 28 days with issue #4's own M30 window (the shared stretch runs Aug 10-30) and 22 days with Monthly Config Watch #1's calendar-August window (shared stretch Aug 10-31), so for those two cohorts that column is decay context rather than pure out-of-sample: the same asymmetry issues #3 and #4 disclosed for their own second columns.

Issue #4's five published cards: home number vs fresh-week replay



Issue #4's five published cards: the number its own issue printed (grey, in-sample on its own window) against the same configuration replayed verbatim on this week (green, out-of-sample). 4 of the five are net-negative on the fresh week; the two M30 cards were measured at home on a four-week window and the three WK cards on a one-week window, so the grey bars are not all the same kind of number. Source: oos_cw4_wk.json + issue #4 open data.

Config	Family	Coins	Home	New WK (n)	New M30 (n)	Verdict
preset 4h4	fable, qwen, deepseek, opus	ETH+SOL	-- live	-\$301.4 (7) * +	+\$41.3 (44)	Faded
preset bc2	gemin, gpt, deepseek	SOL+XRP	-- live	+\$86.3 (4) * +	+\$917.6 (23)	Survived
preset sol	deepseek, fable, opus	SOL [conc]	-- live	-\$263.0 (8) *	+\$1,918.5 (40)	Faded
preset xrp	gemin, gpt, deepseek	XRP [conc]	-- live	-\$387.0 (1) * +	+\$350.2 (11)	Faded
cw4-m301	gpt, deepseek, gemini	SOL+XRP	+\$1,676.0	-\$37.1 (9) * +	+\$1,522.8 (39)	Faded
cw4-m302	fable, gpt, deepseek, gemini	BNB+XRP	+\$1,155.2	-\$255.7 (6) * +	+\$756.1 (28)	Faded
cw4-s1	deepseek	BTC+ETH+ SOL+BNB+XRP	+\$325.0	-\$453.7 (17) +	-\$154.7 (67) +	Faded
cw4-wk1	opus, deepseek, gemini, qwen	BTC+ETH+ SOL+BNB+XRP	+\$399.4	-\$70.0 (16) +	+\$480.7 (64)	Faded
cw4-wk2	fable, deepseek, qwen	BNB+XRP	+\$128.2	+\$60.8 (25)	-\$51.5 (78) +	Faded
cwm1-m1	fable, gemini	SOL+XRP	+\$1,618.4	+\$132.9 (7) *	+\$1,629.9 (34)	Survived
cwm1-m2	fable, gpt, deepseek, gemini, qwen	ETH+SOL+ BNB+XRP	+\$756.6	-\$3.8 (11)	+\$507.5 (40) +	Faded
cwm1-s1	gemin	SOL+XRP	+\$1,721.9	+\$76.6 (8) * +	+\$1,793.9 (44)	Survived
cw3-m301	opus, deepseek, gemini	SOL+XRP	+\$2,137.6	-\$284.7 (20) +	+\$1,239.7 (93)	Faded
cw3-m302	fable, opus, deepseek, gemini, qwen	BTC+ETH	+\$412.3	+\$31.9 (5) *	+\$394.2 (22)	Survived
cw3-s1	gemin	SOL+XRP	+\$1,604.3	-\$209.8 (36) +	+\$213.7 (171)	Faded
cw3-s2	gemin	SOL+XRP	+\$1,449.0	+\$76.6 (8) * +	+\$1,793.9 (44)	Survived
cw3-wk1	opus, deepseek, gemini	SOL+XRP	+\$2,175.1	-\$284.7 (20) +	+\$1,239.7 (93)	Faded
cw3-wk2	opus, deepseek, gemini, grok	BTC+ETH	+\$837.6	-\$362.8 (14) +	-\$633.2 (44) +	Faded
cw2-30d1	opus, gpt, deepseek, gemini, qwen	BTC+ETH+ SOL+BNB+XRP [conc]	+\$744.7	-\$128.3 (5) * +	+\$1,153.9 (43) +	Faded
cw2-30d2	opus, deepseek, gemini	ETH+SOL+ BNB+XRP [conc]	+\$655.1	-\$277.0 (6) * +	+\$753.7 (57) +	Faded
cw2-30d3	opus, deepseek, gemini	BNB+XRP	+\$489.1	-\$239.5 (4) * +	+\$256.9 (31)	Faded

Config	Family	Coins	Home	New WK (n)	New M30 (n)	Verdict
cw2-30ds1	gemini	BNB+XRP	+\$383.8	+\$91.0 (6) *	+\$653.0 (36)	Survived
cw2-7d1	fable, opus, deepseek, gemini	SOL+XRP	+\$299.1	-\$430.7 (16) +	+\$779.8 (89)	Faded
cw2-7d2	gpt, deepseek, gemini, qwen	BNB+XRP	+\$223.7	-\$255.4 (6) * +	+\$500.7 (26)	Faded
cw2-pt1	fable, deepseek, grok [remap]	SOL	+\$698.3	-\$22.1 (11)	+\$993.2 (52)	Faded
cw2-pt2	opus	XRP	+\$543.7	-\$327.5 (4) *	-\$367.0 (15)	Faded
cw1-30d1	fable, deepseek, opus, qwen	ETH+SOL	+\$868.4	-\$244.0 (7) * +	-\$183.2 (45) +	Faded
cw1-30d2	gemini, gpt, deepseek	SOL+XRP+ETH	+\$590.6	-\$133.2 (11) +	+\$705.7 (37)	Faded
cw1-30d3	fable, deepseek, opus, qwen	BTC+ETH+SOL	+\$580.5	-\$246.8 (13) +	-\$254.3 (68) +	Faded
cw1-30ds1	opus	SOL+XRP	+\$442.8	-\$238.4 (23) +	+\$490.2 (99)	Faded
cw1-30ds2	grok [remap]	SOL+XRP	+\$262.9	-\$326.7 (13) +	-\$91.5 (71)	Faded
cw1-30ds3	deepseek	XRP+BNB	+\$119.0	-\$582.0 (14) +	+\$156.8 (115) +	Faded
cw1-7d1	fable, deepseek, opus, gpt	SOL+XRP	+\$439.2	+\$76.4 (32)	+\$1,475.8 (123)	Survived
cw1-7d2	fable, qwen, gemini, opus	XRP+BNB	+\$429.5	-\$135.9 (18) +	-\$141.2 (93) +	Faded
cw1-7d3	gpt, opus	BNB+XRP+SOL	+\$398.6	-\$346.9 (41) +	+\$130.7 (153)	Faded
cw1-7ds1	gemini	XRP+BNB	+\$417.8	+\$23.0 (47)	-\$603.8 (123) +	Faded
cw1-7ds2	gpt	XRP+BNB	+\$366.3	-\$52.3 (35) +	+\$256.0 (127)	Faded
cw1-7ds3	qwen	XRP+BNB	+\$267.0	-\$220.0 (26) +	-\$316.6 (98) +	Faded

n in parentheses; * = fewer than 10 closed trades (insufficient cell); + = non-zero skipped_nofunds (entries the engine could not fund -- the printed economics are then not the economics that ran). Home = the number the card's origin issue printed; the live presets have no home issue. [conc] = stake above this issue's cap for its universe size; [remap] = grok-4.5 replayed on the grok-4.6 lineage. Verdict rule: Survived = net-positive on both fresh windows, Faded = net-negative on at least one. Survivors are tinted green.

Counted: 6 of 34 past cards **Survive** both fresh windows (issue #4 on its two windows: 7 of 26) -- by cohort 0 of 5 issue-4 cards, 2 of 3 Monthly-1 cards, 2 of 6 issue-3 cards, 1 of 8 issue-2 cards, 1 of 12 issue-1 cards. Taken a window at a time: 8 of 34 are positive on the week and 24 of 34 on the four weeks, and the two columns do not nest: a row can clear one and fail the other, so both windows remove cards. The survival curve by generation, each on its own first, second, third and fourth fresh week, reads: issue-1 50.0% -> 91.7% -> 8.3% -> **16.7%** now (2 of 12); issue-2 100.0% -> 50.0% -> **12.5%** now (1 of 8); issue-3 33.3% -> **33.3%** now (2 of 6); issue-4 **20.0%** now (1 of 5); Monthly-1 **66.7%** now (2 of 3). Live presets: 1 of 4 Survive, 1 of 4 positive on the week, 4 of 4 on the month.

The four live presets lead the table by standing rule. On the week: preset 4h4 -\$301.4 on 7 closed, preset bc2 +\$86.3 on 4 closed, preset sol -\$263.0 on 8 closed, preset xrp -\$387.0 on 1 closed -- 1 of the four net-positive. On the four-week window: preset 4h4 +\$41.3 on 44 closed, preset bc2 +\$917.6 on 23 closed, preset sol +\$1,918.5 on 40 closed, preset xrp +\$350.2 on 11 closed -- 4 of four. 1 Survive both. **sol** and **xrp** are marked [conc]: both run a \$900 stake on a single coin, twice the \$450 cap the stake table sets for a 1-coin universe, so their printed economics assume funding this issue's own rule would not grant. Passports for all four are in the open data (presets_recon_wk.json and presets_recon_night.json).

Origin re-runs (drift baseline). Re-simulating the issue-4 cards on their OWN origin windows today lands within -\$186.3..+\$60.3 of the printed issue-4 numbers and the Monthly-1 cards within -\$73.0..+\$40.0 of theirs; 2 of the 8 re-runs reproduce to the cent. The widest deviation is cw4-wk1 (+\$399.4 printed, +\$213.1 today) -- today's engine reads a forecast history that keeps growing, which is the one thing a verbatim replay cannot freeze. Snapshot principle: no past issue is restated; the full drift table is in the open data (oos_walk_forward.drift_baseline).

Honest read. The fresh search and its own back catalogue moved in opposite directions this week. On the fresh week 26 of 34 past cards are net-negative -- issue #4 counted 19 of 26 on its own week -- while the ceiling of the fresh search rose to +\$510.4 from +\$399.4. On the fresh four weeks 24 of 34 are net-positive. Same replay machinery, same frozen knobs; what changed is the window and the size of the board. Three things keep this from being a statement about search. The board is not a fixed sample: it grows every issue and nothing is ever pruned, so a share computed on 34 cards is not comparable with one computed on 26 without saying which generations

joined -- 8 of the 34 rows are new this issue. The month column is not clean evidence for two of the five cohorts: it shares 21 of its 28 days with issue #4's own M30 window and 22 with Monthly Config Watch #1's calendar-August window, so those cards are partly being scored on their own data. And the size of the in-sample number is not what decides the give-back on this board: the row that lost most on the week is cw1-30ds3 (-\$582.0 on the week against +\$119.0 printed at home), while the largest home number on the board, cw3-wk1 at +\$2,175.1, gave back -\$284.7. Four fresh weeks of verdicts now: brutal, the opposite, split by window, and this one. That is a series of weathers, not a strategy.

Calibrate axis: does TP/SL calibration help a config?

Matched pairs -- identical knobs, calibration OFF vs ON -- measured on net PnL. The table is the **whole stage-2 grid** of each window, and the **Both** row pools the two windows' pairs and nothing else. The board-leader slice is a second, separate measurement (12 WK pairs and 14 M30 pairs whose OFF or ON leg reached the top of a board); it is reported in the paragraph below and in the open data, and the two are never pooled together.

Window	Pairs	Calibrate wins	Mean delta	Best delta	Worst delta
WK	1,403	57.8%	+\$136.7	+\$910.0	-\$348.8
M30	1,441	37.8%	-\$65.3	+\$870.3	-\$1,283.2
Both	2,844	47.7%	+\$34.4	+\$910.0	-\$1,283.2

The two windows answer the axis, and the whole-grid reading and the board-leader slice are separate measurements. On the week calibration improves 57.8% of the 1,403 matched pairs (mean +\$136.7, median +\$10.4) and 89.6% of the 385 mature pairs (mean +\$296.5); on the four weeks it improves 37.8% of 1,441 (mean -\$65.3, median -\$10.9) and 51.1% of the 871 mature pairs (mean -\$26.6). Pooled over both windows the axis improves 47.7% of 2,844 pairs with a mean of +\$34.4 -- a number that describes neither window on its own. The board-leader slices are separate: 66.7% of the 12 WK slice pairs improve (mean +\$358.0, median +\$477.4, OFF legs already at a median of +\$15.4), and 0.0% of the 14 M30 slice pairs (mean -\$499.1, OFF legs at +\$1,639.8). Issue #4 measured a 21-pair WK slice (11 improving) and a 15-pair M30 slice (0). Where the axis reached the cards is visible above: 2 of the 2 published WK councils are calibrate=ON cells and 52 of the 62 WK finalists run it, while 0 of the 1 M30 cards are and 27 of the 47 M30 finalists do.

The open-data pack also carries a 400-pair export (calibrate_effect) per collect. Its delta distribution is not the population's (median +\$423.0 against the WK grid's +\$10.4), so no claim in this section is computed from it; it is published for inspection only.

Secondary analysis -- solo-model top

Sidebar to the council narrative. Solo configs run inside the same pipeline and face the same gates: >=2 coins, the maturity gate for the window (>=15 closed on WK, >=20 on M30), one config per model.

Solo top -- WK (Aug 31-Sep 7, 2026)

#	Model	Coins	FH/TF	Net	Net %	maxDD	Trades	WR
#1	gemini	BNB+XRP	4h/1h	+\$178.7	+17.9%	34.4%	44 *	61.4%
#	Ladder	BE	Shifts	Conf filter	Source	Entry	Positional	Stake
#1	50:40,100:60	on	TP -10 / SL +75	>=60 (each)	median	0/0	reenter	\$450

One solo card on this window, not two. Every other row of the mature solo board is either the same model or a single-coin universe, so the one-config-per-model and >=2-coin rules admit exactly one; it runs the 2-coin table stake of \$450 on 44 closed trades with 4 unfunded entries. The two single-coin rows that outrank it on the mature net board (opus on BNB at +\$256.5; opus on BNB at +\$213.8) are excluded by the >=2-coin rule and appear under Per-ticker bests. Source: summary_wk.json top_mature_by_window_branch.

Solo top -- M30 (Aug 10-Sep 7, 2026)

#	Model	Coins	FH/TF	Net	Net %	maxDD	Trades	WR
#1	gemini	SOL+XRP	1d/1d,4h	+\$1,793.9	+179.4%	69.1%	44	72.7%
#	Ladder	BE	Shifts	Conf filter	Source	Entry	Positional	Stake
#1	50:40,100:60	on	TP -10 / SL +75	>=60 (each)	farthest	1/0	reenter	\$450

One solo card, not two on M30. The mature solo board collapses onto gemini (SOL+XRP), so the one-config-per-model rule admits exactly one; the card carries 0 unfunded entries. Source: top_mature_by_window_branch in the night collect.

OPEN IN SANDBOX marketmania.ai/s/cw5-s1

Each link opens this exact frozen window; results visible without sign-in (embedded share signature). The codes are minted at publish time and are NOT live in this draft; every card's full frozen query signature is in the open-data JSON next to its simulation index.

Council vs solo. On the week the best solo sits below the best council (+\$178.7 vs +\$510.4 on the mature board, 44 closed against 66, 34.4% drawdown against 0.0%). On the four weeks the best solo sits above the best council (+\$1,793.9 vs +\$1,717.1, 44 closed against 30, 69.1% against 28.0%). Same in-sample caveats apply to both sides.

Quality boards

Net PnL stays the primary board (continuity with issues #1-#4); four selection views sit on top of it at zero additional simulations. **Calmar-like** = net / max(maxDD, 1.0), positive net only; **WR board** = win rate among mature configs (>=15 closed on WK, >=20 on M30); **Smoothness** = R2 of a linear fit through the daily closed-equity series, mature rows with net > 0 and at least 5 days, ties broken by ulcer index; **Quality composite** = mean rank over net, Calmar-like, win rate and R2, lower is better. Top-3 of each board on each window below; all 15 rows of all eight boards are in the open-data JSON.

Board	Config	Coins	FH/TF	Net	WR	maxDD	CI	R2
WK Calmar	gpt, gemini [cal]	SOL+XRP	4h/4h,1h	+\$510.4	90.9%	0.0%	66	0.899
WK Calmar	gpt, gemini [cal]	SOL+XRP	4h/4h,1h	+\$496.6	89.9%	0.0%	69	0.896
WK Calmar	gpt, gemini [cal]	SOL+XRP	4h/1h	+\$489.2	92.1%	0.0%	63	0.897
WK WR	fable, gpt, gemini [cal]	SOL+XRP	4h/4h	+\$357.0	100.0%	0.0%	32	0.932
WK WR	fable, gpt, gemini [cal]	SOL+XRP	4h/1h	+\$330.8	100.0%	0.0%	33	0.889
WK WR	opus, gemini [cal]	SOL+XRP	4h/4h,1h	+\$304.1	100.0%	0.0%	33	0.867
WK Smoothness	fable, grok [cal]	SOL+XRP	4h/4h,1h	+\$113.8	100.0%	0.0%	17	0.978
WK Smoothness	fable, grok [cal]	SOL+XRP	4h/4h,1h	+\$113.8	100.0%	0.0%	17	0.978
WK Smoothness	fable, grok [cal]	BTC+ETH	4h/4h	+\$145.1	86.4%	0.1%	22	0.967
WK Quality	fable, gpt, gemini [cal]	SOL+XRP	4h/4h	+\$386.3	97.4%	0.0%	38	0.938
WK Quality	fable, gpt, gemini [cal]	SOL+XRP	4h/4h	+\$386.3	97.4%	0.0%	38	0.938
WK Quality	fable, gpt, gemini [cal]	SOL+XRP	4h/4h	+\$357.0	100.0%	0.0%	32	0.932
M30 Calmar	fable, opus, gpt, gemini [cal]	SOL+XRP	4h/4h,1h	+\$1,080.7	96.1%	0.0%	76	0.981
M30 Calmar	fable, opus, gemini [cal]	SOL+XRP	4h/4h,1h	+\$931.3	92.9%	0.4%	85	0.974
M30 Calmar	fable, opus, gpt, gemini [cal]	SOL+XRP	4h/4h,1h	+\$912.6	95.5%	0.0%	67	0.987
M30 WR	fable, deepseek, gemini, qwen [cal]	SOL+XRP	4h/1h	+\$754.4	96.9%	0.5%	65	0.987
M30 WR	fable, deepseek, gemini, qwen [cal]	SOL+XRP	4h/1h	+\$628.7	96.9%	0.4%	65	0.987
M30 WR	fable, deepseek, gemini, qwen [cal]	SOL+XRP	4h/1h	+\$502.9	96.9%	0.3%	65	0.987
M30 Smoothness	fable, gpt, deepseek, gemini, qwen [cal]	SOL	4h/1h	+\$684.5	91.5%	0.8%	47	0.990
M30 Smoothness	fable, opus, gpt, gemini [cal]	SOL+XRP	4h/4h,1h	+\$912.6	95.5%	0.0%	67	0.987
M30 Smoothness	fable, opus, gpt, gemini [cal]	SOL+XRP	4h/4h,1h	+\$608.4	95.5%	0.0%	67	0.987

Board	Config	Coins	FH/TF	Net	WR	maxDD	Cls	R2
M30 Quality	fable, opus, gpt, gemini [cal]	SOL+XRP	4h/4h,1h	+\$1,080.7	96.1%	0.0%	76	0.981
M30 Quality	fable, opus, gpt, gemini [cal]	SOL+XRP	4h/4h,1h	+\$912.6	95.5%	0.0%	67	0.987
M30 Quality	fable, opus, gpt, gemini [cal]	SOL+XRP	4h/4h,1h	+\$900.6	96.1%	0.0%	76	0.981

[cal] = calibrate=ON cell; the board column names the window. Both Calmar boards are degenerate by construction -- their top rows sit at 0.0% (WK) and 0.0% (M30) maxDD, so the score collapses onto net PnL; read them as 'net among configs that barely drew down'. The Smoothness boards are the ones that can leave a window's dominant universe: the WK board's top row is a SOL+XRP config at R2 0.978 on a net of +\$113.8, the M30 board's a SOL config at R2 0.990 on +\$684.5. That trade -- straightness bought with size -- is exactly what the board exists to expose. The M30 boards are drawn from a mature pool of 3,252 configs against 1,674 on WK, because four weeks clear a >=20-trade gate that one week does not.

Stability, and what it is across this issue. The board is defined as **mature net top-40 on BOTH windows** of the run that computes it. The run that produced it here is the night collect, whose two windows are **WK and M30** -- the same two windows this issue prints, computed in a second execution of the WK plan -- so this issue's board reads **stable across WK and M30. 0 configurations qualify**. An empty board is a result, not a missing file: the night collect ran all 8 of its phases with 0 errors and wrote the key as an empty list. The two windows disagree about which configurations are good enough to be on a board at all: 9 of the 20 rows exported from its WK board are BNB; all 20 of the 20 rows exported from its M30 board are SOL+XRP, and no row exported from one window's mature board appears on the other's. The count is not comparable with issue #4's 25 (its board was measured across M30 and the calendar month); the last board measured across WK and M30 was issue #3's, at 21. The board is empty, so there is no row list in the open data either -- `stability.rows` is `[]` and `stability.n\_configs` is 0.

Hand-tuned baseline vs the grid

Standing section, second issue. One hand-tuned configuration is treated as the reference the search has to beat -- not on net PnL, where any large search wins by construction, but on the three properties a configuration is actually tuned for: **win rate, shallow drawdown and a straight equity line**. The reference (referred to below as the **hand-tuned anchor**) is the live preset 1d_3_llm_BC2_v02, a 3-model 1d council on SOL+XRP at the \$450 two-coin stake; it is distinct from the pre-search hand-tuned set A/B/C carried in Baselines below. Its replay row is in the walk-forward table above; here it is the yardstick.

Beat rule (frozen with the criteria): a config beats the anchor only if it does so on all three at once -- win rate higher, maxDD lower and smoothness R2 higher -- among mature, non-concentrated search rows. Net PnL is not part of the rule; it is reported next to the winners so the cost of the improvement is visible.

The anchor on each window

Window	Config	Coins	FH/TF	Net	WR	maxDD	R2	Ulcer	Cls
WK	Hand-tuned anchor (1d_3_llm_BC2_v02)	SOL+XRP	1d/1d	+\$86.3	75.0%	19.4%	0.039	9.55	4
M30	Hand-tuned anchor (1d_3_llm_BC2_v02)	SOL+XRP	1d/1d	+\$917.6	78.3%	44.0%	0.691	8.18	23

315 of the 1,674 mature WK configs (18.8%) clear all three bars at once, and 370 of the 3,252 mature M30 configs (11.4%). The anchor's own two windows are different rows: on the week +\$86.3 net at 75.0% win rate with a 19.4% drawdown, an R2 of 0.039 and 4 closed trades; on the four weeks +\$917.6 at 78.3%, a 44.0% drawdown, an R2 of 0.691 and 23 closed. The week row is thin against the >=15 gate the search rows must pass (the anchor is exempt from that gate by construction: it is the reference, not a candidate). Read against net PnL, all 10 of the 10 best WK winners by net also beat it on net PnL; all 10 of the 10 best M30 winners by net also beat it on net PnL. The share itself moves with the window (18.8% of the WK field, 11.4% of the M30 field), so read it as one field at a time, not as a property of the anchor.

Top-10 by net among the 315 WK configs that beat the anchor on win rate, maxDD and smoothness at once

#	Config	Coins	FH/TF	Stake	Net	WR	maxDD	R2
#1	gpt, gemini [cal]	SOL+XRP	4h/4h,1h	\$450	+\$510.4	90.9%	0.0%	0.899
#2	gpt, gemini [cal]	SOL+XRP	4h/4h,1h	\$450	+\$496.6	89.9%	0.0%	0.896
#3	gpt, gemini [cal]	SOL+XRP	4h/1h	\$450	+\$489.2	92.1%	0.0%	0.897
#4	gpt, gemini [cal]	SOL+XRP	4h/1h	\$450	+\$489.2	92.1%	0.0%	0.897
#5	gpt, gemini [cal]	SOL+XRP	4h/4h,1h	\$450	+\$465.9	96.1%	0.0%	0.898
#6	gpt, gemini [cal]	SOL+XRP	4h/4h,1h	\$450	+\$438.4	95.1%	0.0%	0.882
#7	gpt, gemini [cal]	SOL+XRP	4h/4h,1h	\$375	+\$425.3	90.9%	0.0%	0.899
#8	gpt, gemini [cal]	SOL+XRP	4h/4h	\$450	+\$415.3	87.0%	0.0%	0.945
#9	gpt, gemini [cal]	SOL+XRP	4h/4h	\$450	+\$415.3	87.0%	0.0%	0.945
#10	gpt, gemini [cal]	SOL+XRP	4h/4h,1h	\$375	+\$413.8	89.9%	0.0%	0.896

Top-5 by net among the 370 M30 configs that beat the anchor on win rate, maxDD and smoothness at once

#	Config	Coins	FH/TF	Stake	Net	WR	maxDD	R2
#1	fable, opus, gpt, gemini	SOL+XRP	1d/4h	\$450	+\$1,717.1	80.0%	28.0%	0.871
#2	fable, opus, gpt, gemini	SOL+XRP	1d/4h	\$450	+\$1,717.1	80.0%	28.0%	0.871
#3	fable, opus, gemini	SOL+XRP	1d/4h	\$450	+\$1,497.0	78.6%	28.0%	0.865
#4	fable, opus, gemini	SOL+XRP	1d/4h	\$450	+\$1,497.0	78.6%	28.0%	0.865
#5	fable, opus, gpt, gemini, grok	SOL+XRP	1d/1d	\$450	+\$1,491.1	80.0%	41.3%	0.802

Anchor passport (verbatim, as stored): members=gemini-3.1-pro,gpt-5.6-sol,deepseek-v4-pro tokens=SOL,XRP fh=1d tf=1d conf_mode=each min_conf=60 max_sideways=2 max_diff_side=0 tp_sl_source=nearest trade_mode=position same_side=update opposite=reverse no_signal=hold time_stop=3x steps=40:80,100:20 steps_on=1 be=1 tp_shift=30 sl_shift=50 deposit=1000 lev=10 refill=0 stake=450 calibrate=- variant=v0

Per-ticker bests

Which coin was extractable this window, and by what. Single-coin results are excluded from the council cards by rule (idiosyncratic-coin risk) and reported here, never ranked against the cards. The dedicated per-coin branch contributed 580 + 1,115 sims -- a compacted grid, one best per coin per window; the branch tag in brackets says which kind won.

WK (Aug 31-Sep 7, 2026)

Coin	Best config (branch)	FH/TF	Net (trades)
BTC	gemini, grok [council]	1d/1d	+\$138.9 (3) *
ETH	opus, deepseek [council]	1d/1d	+\$158.3 (4) *
SOL	gpt, gemini [council]	4h/4h,1h	+\$314.0 (31)
BNB	opus, grok [council]	4h/1h	+\$297.2 (17)
XRP	gpt, gemini [council]	4h/1h	+\$215.4 (33)

M30 (Aug 10-Sep 7, 2026)

Coin	Best config (branch)	FH/TF	Net (trades)
BTC	fable, opus, gpt, gemini [council]	1d/4h	+\$506.7 (17)
ETH	deepseek, gemini, qwen [council]	1d/1d	+\$550.4 (15)
SOL	gemini [solo]	1d/1d	+\$1,256.8 (21)
BNB	fable, gemini, qwen [council]	1d/1d	+\$635.3 (12)
XRP	opus, deepseek, gemini [council]	4h/1h	+\$855.2 (72)

n in parentheses; * = fewer than 10 closed trades (thin -- direction only). 3 of the ten winners run calibrate ON; 5 of the five WK winners and 4 of the five M30 winners are councils.

The reads. On the week SOL leads at +\$314.0 and 2 of the five rows are thin -- 3 to 33 closed trades against a 15-trade gate. On the four weeks no row is thin: 12 to 72 closed, SOL leading at +\$1,256.8, and the same coin leads both windows. The 1d horizon wins 2 of the five WK coins and 4 of the five M30 coins. **Winner's curse applies to this whole table:** each row is the maximum of a per-coin grid, biased high by selection alone, and none has an out-of-sample read until next issue.

Baselines

Three independent reference points, none of them search output, re-simulated fresh on both of this issue's frozen windows inside the same two collects (engine 1.1): the engine solo default for each of the 7 models, the default council-of-7, and the pre-search hand-tuned set A/B/C carried since issue #1. The tables show the universe-aware **table-stake** run -- this issue's canonical mode (\$200 for the 5-coin defaults; A and B already at their \$300 3-coin cap; C moves \$400 -> \$450 on 2 coins) -- with the best and worst of the seven solo defaults on each window; all eleven rows per window and both stake modes are in the open data.

WK (Aug 31-Sep 7, 2026)

Baseline	Net PnL	Trades	WR	maxDD
Best solo default (gemini-3.1-pro)	-\$40.1 *	113	52%	4.3%
Worst solo default (deepseek-v4-pro)	-\$86.2 *	328	56%	8.6%
Council-of-7 default	-\$40.7 *	131	53%	4.1%
Hand-tuned config A (pre-search)	+\$155.4	4	100%	0.0%
Hand-tuned config B (pre-search)	-\$251.6 *	44	52%	28.2%
Hand-tuned config C (pre-search)	-\$244.0 *	7	43%	29.8%

M30 (Aug 10-Sep 7, 2026)

Baseline	Net PnL	Trades	WR	maxDD
Best solo default (gemini-3.1-pro)	-\$48.2 *	470	56%	10.1%
Worst solo default (deepseek-v4-pro)	-\$183.4 *	1069	56%	18.3%
Council-of-7 default	-\$77.8 *	439	57%	8.2%
Hand-tuned config A (pre-search)	+\$498.7	24	79%	50.1%
Hand-tuned config B (pre-search)	+\$215.5	164	64%	78.2%
Hand-tuned config C (pre-search)	-\$183.2 *	45	58%	69.0%

* = non-zero skipped_nofunds at table stake. On the 5-coin defaults those counts are the stake table doing its job: at \$200 a \$1,000 deposit funds at most five concurrent positions while the defaults fire hundreds of entries per window (107 to 377 skipped each on the week, 416 to 1,169 on the four weeks). The other five solo defaults run between the two printed rows on each window. **The conclusions do not change between modes:** on WK they do so for every row, and on M30 for every row but one, printed as it stands: solo default gemini-3.1-pro is +\$18.3 at the config's own stake and -\$48.2 at the table stake. 2 of the three hand-tuned configs are already at their own table cap, so their two modes are identical on both windows. Source: baselines_wk.json (22 sims) and baselines_night.json (44 sims, of which 22 are the M30 half), inside this issue's two collects.

Every engine default lost money on both windows; the hand-tuned set did not. On WK the seven solo defaults run -\$40.1 to -\$86.2 at table stake, the council-of-7 -\$40.7, and the three hand-tuned configs A +\$155.4, B -\$251.6, C -\$244.0 -- 1 of the three positive. On M30 every default is negative again (-\$48.2 to -\$183.4 solo,

council-of-7 -\$77.8) and the three hand-tuned configs A +\$498.7, B +\$215.5, C -\$183.2 -- 2 of the three positive. The upper half of the standing ordering (search > hand-tuned > defaults) holds on both windows, in-sample as ever -- the fresh search makes +\$510.4 on the week against a best baseline of +\$155.4, and +\$1,717.1 on the four weeks against +\$498.7. The live presets are a separate row of evidence: 1 of the four is net-positive on the week and 4 of four on the month (walk-forward table above), and 2 of them run stakes above this issue's cap.

Live presets & preset neighbourhoods

Four hand-tuned configurations trade live on the platform (not search output, distinct from the pre-search A/B/C set): preset bc2 = '1d_3_llm_BC2_v02'; preset 4h4 = '4h_4_llm_v02'; preset sol = 'Sol'; preset xrp = 'XRP'. Their replay rows lead the walk-forward table; passports and window detail are in the open data (presets_recon_wk.json and presets_recon_night.json). The neighbourhood grid asks a narrower question: does a one-knob neighbour beat the preset as configured? Two of the four have a local grid this issue (53 cells per window); the two single-coin presets do not (0 cells -- the local grid is not built for them, so no neighbour claim is made about either).

Live preset	Window	As configured	Best neighbour in local grid	Cells	Verdict
preset bc2	WK	+\$86.3	+\$279.8 (1d/1d, SOL, 3 cls)	34	neighbour +\$193.5
preset bc2	M30	+\$917.6	+\$1,159.7 (1d/1d, SOL+XRP, 22 cls)	34	neighbour +\$242.1
preset 4h4	WK	-\$301.4	+\$95.4 (4h/1h, ETH+SOL, 16 cls)	19	neighbour +\$396.8
preset 4h4	M30	+\$41.3	+\$637.2 (1d/1d,4h, ETH+SOL, 9 cls)	19	neighbour +\$595.9

In **4 of the 4** preset-window cells with a grid, a one-knob neighbour beat the configuration as it is actually running. The largest single edit is preset 4h4's 4h/1h -> 1d/1d,4h horizon move on M30, which turns +\$41.3 into +\$637.2. The presets are not at a local optimum -- a testable, low-risk edit, unlike adopting a search champion wholesale. Cell counts are small (34 and 19 per window) and every neighbour is an in-sample maximum of its own little grid.

Patterns: what the finalists look like

Knob modes across each window's unique finalists (mature net top-10 per branch plus all four quality boards, deduplicated by full signature; 62 unique configs on WK, 52 of them councils; 47 on M30, 37 councils).

Knob	WK finalists mode	M30 finalists mode
Forecast horizon	4h (62/62)	4h (25/47)
Ladder shape (steps)	50:40,100:60 (62/62)	50:40,100:60 (46/47)
Break-even stop (BE)	on (62/62)	on (47/47)
Min confidence	60 (62/62)	60 (47/47)
SL shift	75 (62/62)	75 (46/47)
TP shift	-10 (62/62)	-10 (46/47)
TP/SL source	farthest (57/62)	farthest (41/47)
Calibrate ON	52/62	27/47
Universe	SOL+XRP (40/62)	SOL+XRP (46/47)

The mechanical core is the same on both windows and unchanged for the fifth issue running: 50:40,100:60 ladder, break-even on, conf 60, SL +75 / TP -10, farthest source. The horizon is **4h** on WK (62/62) and **4h** on M30 (25/47). **Calibrate=ON** runs in 52 of 62 WK finalists against 27 of 47 on M30, where issue #4 measured 26/40 on its week and 35/62 on its month. The universe: SOL+XRP takes 40 of the 62 WK finalists while SOL+XRP takes 46 of the 47 M30 finalists. Membership concentrates too -- gemini leads the WK council finalists (40 of 52) and gemini the M30 ones (37 of 37). Standing caveats: refinement seeds from the same grid winners, so part of the convergence is search-design echo; and a knob core that produced 7-of-26 survival one issue ago and 6-of-34 this issue is describing the weather, not a recommendation.

Liquidation note

Liquidation is modeled by the engine at the fixed 10x leverage used throughout. Across the published cards, stop-loss placement stays inside the distance that would approach the liquidation threshold; maxDD is printed on every config so realized risk is visible directly. This issue the published cards draw down 0.0% and 0.1% on the week and 28.0% on the month, and the replay table is deeper: 8 replayed rows drew down more than 40% on the week and 29 on the month, 2 and 23 of them past 55%. A high headline net and a survivable path are different claims, and so are a shallow in-sample drawdown and a shallow one next week.

Watch amendments (this issue)

Amendment 3 (standing) -- smoothness and a quality composite. Every simulation stores its daily closed-equity series and the three numbers read off it (n_days, smooth_r2, ulcer); the smoothness board ranks mature rows with net > 0 and at least 5 days by R2 with ulcer as tie-break, and the quality composite ranks by the mean of the net, Calmar-like, win-rate and R2 ranks. Both still cost zero additional simulations and the headline board is still net PnL.

Amendment 4 (standing) -- the monthly split. The weekly Config Watch keeps two windows, the calendar week (WK) and the four calendar weeks that end at the same edge (M30, Aug 10-Sep 7, 2026); the calendar-month board belongs to the Monthly Config Watch line and no calendar-month number is printed in a weekly issue. This issue's night collect computes no calendar-month window at all: the next monthly issue covers September and is produced after 01.10.

Amendment 5 -- the standing OOS board grows with every issue. Every card this series publishes joins the walk-forward table and is never removed: this issue replays **34 cards** -- issue #4's 5 and Monthly Config Watch #1's 3 join issues #1, #2 and #3's 26 -- plus the four live presets, on both fresh windows. **Monthly cards are judged by the weekly rule** (net-positive on both of this issue's fresh windows), the same rule every other row is judged by; their own calendar-month origin window is the Home column and nothing else. One consequence for the counts: a survival share is computed on a board that changes size every issue, so it is reported with the cohort breakdown beside it and never as a trend on its own. **Stability** is defined inside the run that computes it: this issue's night collect computes WK and M30, so the board reads across WK and M30 -- the issue-#3 definition -- where issue #4's read across M30 and the calendar month. It comes back with **0 configurations** this issue: an empty board is a result and is printed as one, with no row list and no top-three table.

All three amendments are provisional and scoped to this series. **Formalization is slated for methodology v1.2;** until then this issue runs on v1.1 plus the Config Watch amendments approved 12.08 (hash e66c7e8c864a2233).

Limitations

- **Two windows, one regime, and they overlap.** M30 (Aug 10-Sep 7, 2026) contains WK (Aug 31-Sep 7, 2026) -- 7 of the week's 7 days are inside it -- and it shares 21 of its 28 days with issue #4's own M30 window and 22 with Monthly Config Watch #1's calendar-August window. The two columns of this issue are therefore not two independent tests: agreement between them is partly arithmetic, and for the issue-4 and Monthly-1 cohorts the M30 replay column is decay context rather than pure out-of-sample. Do not read them as a cross-validation.
- **The two windows come from two collects, and the two runs of the same week do not agree.** WK is the weekly run (generated 2026-09-08T12:32:32Z), M30 the night run (2026-09-09T04:20:39Z -- boards computed at 04:20:39Z, files written after the baselines pass), which recomputed the same WK week in the same pass on its own grid: the night run's WK ceiling is +\$417.4 (fable, gemini, qwen on BNB) against the weekly collect's +\$510.4. Every WK number this issue prints is the weekly collect's; the night run's WK half is used only for the cross-window stability board. Same engine, same criteria, same collect code, but not the same execution -- a number is comparable across the two windows only as far as that is.
- **Stability is defined inside the run that computes it, and this issue's board is empty.** The board comes from the night run, whose two windows are WK and M30, so it reads across those two -- and 0 configurations are in the mature net top-40 of both. There is therefore no cross-window stability evidence in this issue at all, and no claim is made from its absence: one empty board is one observation. Issue #3's board (the last across the same pair) held 21; issue #4's 25 is not comparable at all, having been measured across M30 and the calendar month.
- **Model lineage is a splice.** The grok axis runs as grok-4.6 (incl. 4.5-era) -- one lineage across a mid-series model cutover dated 2026-08-24 in the search plan both collects ran to. This issue's WK window (Aug 31-Sep 7, 2026) lies entirely after that cutover; the M30 window (Aug 10-Sep 7, 2026) contains it. Issue-1 and issue-2 configs that named grok-4.5 are replayed on the 4.6 lineage ([remap] rows: cw2-pt1, cw1-30ds2), and a remapped replay is not the same simulation the origin issue ran.

- **Concentrated rows are published, not hidden.** 4 rows run stakes above this issue's cap for their universe size and are marked [conc] (preset_sol, preset_xrp, cw2-30d1, cw2-30d2); their printed economics assume funding the current rule would not grant. They stay in the table because removing them would flatter the preset row.
- **The stake rule changes what runs, not only its size.** Capping the stake changes WHICH entries get funded: the seven table-stake solo defaults skip 107 to 377 entries each, and 28 replay rows carry unfunded entries (up to 16 on cw1-30ds3). Where skipped_nofunds is non-zero the printed economics are not the economics that ran; the rows are flagged, never silently pooled.
- **Snapshot principle.** This report is a frozen snapshot taken at the generated-at timestamps; numbers are not updated retroactively and past issues are not restated. The drift baseline exists precisely because today's engine and a longer forecast history do not reproduce every past number (cw4-wk1: +\$399.4 printed, +\$213.1 today).
- **Research-to-date counter is pinned.** The counter below is pinned from wave 3 onward and is read at the Sep 7, 2026 16:00 UTC cut-off; issue #4 printed 38,385 resolved and 18 published reports against 42,905 and 26 here. The model line keeps the definition issue #4 introduced (7 tracked, current line-up; earlier versions folded into their successors' lineage), so the two issues' model counts are comparable and issue #3's is not.
- **No calendar month in this issue.** The night collect computes M30 and WK and nothing else, so no claim about a calendar month is made anywhere here. Monthly Config Watch #2 covers September and is produced after 01.10; Monthly Config Watch #1's three cards appear in this issue only as replay rows, judged by the weekly rule.
- **Multiple testing.** 16,398 simulations across the two collects (5,435 weekly + 10,963 night); at this scale some winners are expected from chance alone. Antidotes: axes frozen before launch, the walk-forward table, independent baselines, in-sample labeling. No formal correction yet.
- **Winner's curse and thin cells.** Every card, board row and per-ticker best is the maximum of a search, biased high by selection alone. On the week the per-ticker winners closed 3-33 trades and 17 of the 38 replay rows are under 10 closed; on the four weeks no replay row is thin and no per-ticker winner is. Flagged with *, reported for completeness.
- Research output, not financial advice.

Counters & lineage

Weekly collect (WK window)

Stage	Sims	Stage	Sims
Scan (s1)	476	Dedicated per-ticker (s6t+s6tc)	580
Systematic grid (s2)	4,080	OOS replays (s4cw4+s4cw4o+s4cwm1+s4cwm1o+s4cw3+s4cw3o+s4cw2+s4cw2o+s4cw1+s4pre)	60
Refinement: ladder + stake sweep (s2b+s3b)	116	Total	5,435
Preset neighbourhood grid (s2p)	53	Errors / zero-trade	0 / 996
Per-ticker probes (s5t)	70	Engine	1.1

Night collect (M30 and a second execution of the WK week; this issue publishes M30 from it, and its WK half feeds the cross-window stability board only)

Stage	Sims	Stage	Sims
Scan (s1)	952	Dedicated per-ticker (s6t+s6tc)	1,115
Systematic grid (s2)	8,352	OOS replays (s4cw4+s4cw4o+s4cwm1+s4cwm1o+s4cw3+s4cw3o+s4cw2+s4cw2o+s4cw1+s4pre)	98
Refinement: ladder + stake sweep (s2b+s3b)	230	Total	10,963
Preset neighbourhood grid (s2p)	106	Errors / zero-trade	0 / 1,711
Per-ticker probes (s5t)	110	Engine	1.1

Scale note (vs issue #4). Issue #4 ran 16,447 sims over three windows in two collects; issue #5 runs 16,398 over two windows in two -- 5,435 in the weekly collect (WK) and 10,963 in the night collect (M30 and the same WK week together). The per-window grid is close to unchanged: 4,080 WK cells and 4,176 M30 cells here against 4,208 per window there. The per-coin branch is 580 and 1,115 against 585. The replay branch grew in rows (38 configs against 30) and in sims (60 and 98 against 44 and 74), because two more cohorts joined and every row is replayed on every window of its collect. The two selection boards still cost zero simulations.

No extras pass in either collect. Equity series, the grid histograms (0 extra sims), the baselines and the calibrate packs are produced by the same two collects that wrote the boards. The baseline runs are verification, outside both search totals: 22 sims on WK (11 configs x 2 stake modes) and 44 in the night collect, of which the 22 M30 rows belong to this issue. 0 equity replays were needed.

Generated at: 2026-09-08T12:32:32Z (weekly collect) and 2026-09-09T04:20:39Z (night collect) · Methodology v1.1 + Config Watch amendments (approved 12.08), hash e66c7e8c864a2233.

Snapshot principle: this report is a frozen snapshot of the search taken at the generated-at timestamps above; numbers are not updated retroactively. Each issue re-runs the pipeline fresh over that issue's windows.

Research to date

RESEARCH SNAPSHOT

THIS REPORT -- search effort: 16,398 simulations (0 errors) across two frozen collects, of which 158 walk-forward replays and origin re-runs · 66 verification sims (baseline runs inside the same collects; 0 equity replays needed)

MARKETMANIA RESEARCH TO DATE (as of cutoff): 42,905 directional forecasts resolved since Jul 11 (as of Sep 7, 2026 16:00 UTC cutoff) · 7 models tracked (current line-up; earlier versions folded into their successors' lineage) · 5 assets · 5 horizons · hourly · 26 published reports

Counter pinned from wave 3 onward and read at the Sep 7, 2026 16:00 UTC cut-off: issue #4 printed 38,385 resolved and 18 published reports. The model line keeps the definition issue #4 introduced (7 tracked, current line-up; earlier versions folded into their successors' lineage), so the two issues' model counts are comparable. Config Watch #5 is the 26th published report.

Issue #5. The standing walk-forward board reaches 34 cards -- issue #4's 5 and Monthly Config Watch #1's 3 join it, and monthly cards are judged by the weekly rule (Amendment 5). Two fresh windows again, the calendar week and the four calendar weeks ending at the same edge, both computed in the night collect as well as the week: the table keeps 6 of 34 past cards where the week alone keeps 8 and the month alone 24. The cross-window stability board is measured on this issue's own two windows (WK and M30) and comes back with 0 configurations. Formalization of the amendments is still slated for methodology v1.2.

What we're testing next

- **The first fresh week of this issue's own cards.** Config Watch #6 replays the 4 cards published here that carry a short link (both WK councils, the WK solo card, the M30 council; the M30 solo card has no link and is not replayed, as issue #4's was not) on its own fresh week, alongside the 34 already on the board. The number to check against: 20.0% of issue #4's 5 linked cards were net-positive on their first fresh week, which is this issue's week.
- **Whether the WK-and-M30 stability board fills again.** 0 configurations are in the mature net top-40 of both windows of this issue's night run, against 21 on issue #3's board across the same pair. Config Watch #6's night run computes the same pair one week on, so the question is whether the board is empty twice running -- and it is the first time this board can be compared with itself, because issue #4's was built across different windows.
- **Whether the calibrate share holds for a third window in a row.** Calibration improves 57.8% of the 1,403 matched WK pairs and 37.8% of the 1,441 M30 pairs this issue, against 61.0% and 29.1% in issue #4. The test is the same two measurements on the next pair of windows -- the whole grid and the board-leader slice, reported separately and never pooled.

Related research

Consensus Watch #5 -- <https://marketmania.ai/research/reports/consensus-watch-2026-08-31.pdf>

Weekly Calibration #5 -- <https://marketmania.ai/research/reports/weekly-calibration-2026-08-31.pdf>

Weekly Model Watch #5 -- <https://marketmania.ai/research/reports/model-watch-2026-08-31.pdf>

Config Watch #4 -- <https://marketmania.ai/research/reports/config-watch-2026-08-26.pdf>

Config Watch #3 -- <https://marketmania.ai/research/reports/config-watch-2026-08-19.pdf>

Monthly Config Watch #1 -- <https://marketmania.ai/research/reports/config-watch-monthly-2026-08.pdf>

The three weekly reports publish together as one issue each week; Config Watch follows its own cycle. Market-regime figures, where this issue refers to them, come from that wave and are not re-tabulated here. Monthly Config Watch #1 is listed because its three cards are replayed in the walk-forward table above.

CITE THIS REPORT

```
@techreport{mm_configwatch_2026w37,  
  title      = {Config Watch #5},  
  author     = {{MarketMania Research}},  
  institution = {MarketMania},  
  year       = {2026},  
  month      = sep,  
  day        = {9},  
  type       = {Weekly Research Report},  
  series     = {Config Watch},  
  number     = {5},  
  note       = {Methodology v1.1 + Config Watch amendments, hash e66c7e8c864a2233},  
  url        = {https://marketmania.ai/research/reports/config-watch-2026-09-02.pdf}  
}
```